THE UNIVERSITY OF ALBERTA

A COMPARISON OF A GROUP AND AN INDIVIDUAL INTELLIGENCE TEST USING STUDENTS WITH ADJUSTMENT DIFFICULTIES

BY

© CHRISTINE M. PERNAROWSKI

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF EDUCATION

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

EDMONTON, ALBERTA

FALL, 1972

THE UNIVERSITY OF ALBERTA

A COMPARISON OF A GROUP AND AN INDIVIDUAL

INTELLIGENCE TEST

USING STUDENTS WITH ADJUSTMENT DIFFICULTIES

BY

 CHRISTINE M. PIENKOWSKI

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF EDUCATION

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

EDMONTON, ALBERTA

FALL, 1972

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled A Comparison of a Group and An Individual Intelligence Test Using Students With Adjustment Difficulties, submitted by Christine M. Pernarowski in partial fulfillment of the requirements for the degree of Master of Education.

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled A Comparison of a Group and An Individual Intelligence Test Using Students With Adjustment Difficulties, submitted by Christine M. Fernandez in partial fulfillment of the requirements for the degree of Master of Education.

ABSTRACT

The purpose of this study was to compare a group-administered intelligence test with an individually-administered intelligence test using a random sample of students with adjustment difficulties (adjustment sample) and a random sample of all students (control sample). It was expected that the adjustment sample would score higher in the individual test than on the group test.

The samples were selected from the Grade four, five and six students of the Biggin Hill School District, Medley, Alberta. The adjustment sample included thirty students from those in Grades four, five and six who had been seen by a counselor at least four times during the 1971-72 term. The control sample included thirty students from the total population in Grades four, five and six.

The "Wechsler Intelligence Scale for Children" (WISC) and the "Otis-Lennon Mental Ability Test" (Otis-Lennon) were administered to all students in both samples. A comparison of the results indicated that there was no difference in the WISC and Otis-Lennon scores for the control sample. The adjustment sample, however, scored higher on the WISC than on the Otis-Lennon. An analysis of the data determined that although the difference in the WISC and Otis-Lennon scores for the adjustment sample was approaching significance, it was not significant at the .05 level. Also, the Otis-

Lennon correlated as highly with the WISC (Full Scale) for the adjustment sample as it did for the control sample. The WISC Verbal scores, however, did not correlate as highly with the Otis-Lennon for the adjustment sample as they did for the control sample.

Because the difference in the WISC (Full Scale) and Otis-Lennon deviation intelligence quotients (DIQs) of the adjustment sample approached significance and the WISC Verbal scores correlated more highly with the Otis-Lennon for the control sample than they did for the adjustment sample, the author suggested that an individual test be used for students experiencing adjustment difficulties.

ACKNOWLEDGEMENTS

The writer wishes to acknowledge the assistance of those who contributed to the completion of this thesis.

Sincere thanks are due to Dr. D. Sawatzky for his patience, helpful observations and advice; to Dr. J. Paterson for his constructive suggestions and encouragement; and to Dr. E. Romaniuk for his thoughtful consideration and assistance with the statistical design.

Special thanks are also extended to the administrators, teachers and students of the Biggin Hill School District who assisted or participated in this study. To Mrs. Donna Stephens thanks are expressed for typing so capably.

The writer also wishes to express her appreciation to her parents and sisters for their interest and support, especially to Mrs. Louise Cundict who so willingly discussed questions related to the study.

Finally, special thanks are extended to her husband, Ronald, for his continual patience, understanding and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Standardized tests have been used increasingly in recent years to make educational and psychological assessments. Generally tests are administered to obtain information which will help in decision making. Although they supply useful data, the results must be used cautiously because tests measure only test-taking behavior. Brown (1970) said that:

> Psychological tests are by no means perfect and no reputable psychologist will say they are. Yet even though tests do make mistakes in individual cases, there is a body of evidence that shows that in a variety of situations tests do a better job than other available educational methods. This point is often overlooked. In many decisions for practical purposes the question is not a black-white, accurate-inaccurate distinction, but rather which procedure makes the most correct decisions (p. 2).

There has been much discussion regarding the word "intelligence" and the use of intelligence tests. Regardless of the difficulties we have in defining "intelligence" and in deciding whether or not we should use intelligence tests, there are numerous intelligence tests being administered in schools.

Tests used in the school setting are administered to groups and individuals. Because the group-administered tests are relatively easy and rapid to administer they are given more frequently than the individually-administered

tests. The group-administered tests have other advantages too, but the individually-administered tests usually are considered more accurate because the examiner has a greater opportunity to establish rapport and obtain the co-operation of the examinee.

In order to determine if the results of group tests administered to students experiencing adjustment difficulties are used more frequently by the school personnel than the results of tests administered to the average students, a survey was conducted in Northeastern Alberta. The results of this telephone survey, which was conducted in two school systems, indicated that intelligence scores of children who are experiencing adjustment problems seem to be used more frequently by teachers, counselors and administrators than are the intelligence scores of other students. If the intelligence scores of students with adjustment problems are the scores used more frequently, then the intelligence tests which are administered should supply educators with accurate information about these students. Generally, the intelligence tests given in schools are group-administered tests.

Publishers of group-administered intelligence tests have compared them to individually-administered tests and have often obtained favorable results. For example, Otis and

Lennon (1969) compared the "Otis-Lennon Mental Ability Test" to the "Stanford Binet Intelligence Scale" and found the correlation to be .60. The Stanford Binet had been administered twenty-one months before the Otis-Lennon so the authors, Otis and Lennon (1969), concluded that:

> The correlation obtained, especially in view of the time interval, indicates that similar attributes are being measured by both tests (p. 37).

The samples used for the comparisons of group and individually-administered tests are usually random samples of all school children. Since counselors, teachers, and administrators appear to use the intelligence scores of children with adjustment difficulties more than they use the intelligence scores of other children, it seemed worthwhile to compare the group-administered test to the individually-administered test using a random sample of youngsters experiencing adjustment difficulties. The author's purpose was to determine whether the correlation of an individual and a group test will be as high using a random sample of youngsters with adjustment difficulties, as it would using a random sample of all youngsters. The tests used for this study were the "Otis-Lennon Mental Ability Test" (Otis-Lennon) and the "Wechsler Intelligence Scale for Children" (WISC). These tests were chosen because they are known as two of the finest intelligence tests of their kind (Adams, 1964; Anastasi,

1968; Brown, 1970; DuBois, 1970; Glasser and Zimmerman, 1968; Grotelueschen, 1969; Horst, 1966; Smith, 1969; Tyler, 1963).

Although most children are given intelligence tests in schools, it appears that the results used most frequently are the scores of children with adjustment difficulties. Very often the intelligence test scores available for these youngsters are scores obtained from group-administered intelligence tests. Therefore, the purpose of this study was to compare a group-administered test (Otis-Lennon) with an individually-administered test (WISC). The tests were administered to two random student samples. One group was a random sample of children who were experiencing adjustment problems (adjustment sample) and the other group was a random sample of all children (control sample). It was expected that the correlation between the two tests would be lower if the children were experiencing adjustment difficulties.

In summary, the purpose of this study was to determine to what degree the Otis-Lennon correlates with the WISC when administered to students with adjustment difficulties.

CHAPTER II

RELATED LITERATURE AND HYPOTHESES

RELATED LITERATURE

Background

Measurement has been defined as the assignment of numerals to objects or events according to rules (Stevens, 1951). Brown (1970) explained that what we actually measure is the properties and characteristics of these objects and events. More precisely, we measure indicants of the properties and characteristics of objects and events. Test scores are the indicants and from these we make inferences about the characteristics of the object. Some educators over estimate the accuracy of test scores while others would contend that tests are useless and wrong decisions will be made if test scores are used. Both views are extreme. Test scores do have limitations but they can provide useful information.

Brown (1970) stated that no test is so comprehensive that it measures every possible aspect of a given behavior domain. There are basically two problems: (1) choosing a representative sample of the universe of possible behaviors, and (2) deciding whether a person would obtain the same score using a different sample of the same domain. If a test is representative of the universe of possible behaviors

which it claims to measure, it is considered a valid test. Likewise, if the examinees obtain the same scores when tested with a different sample from the same domain, the test is considered reliable. The more valid and reliable a test is, the more accurate the results will be and as a result better decisions will be made. In the school setting the results are often used for selection, placement, diagnosis, hypothesis testing, and evaluation. Jenson (1969) stressed the importance of accurate measurement when he stated:

> Whether we like it or not, the educational system today is one of society's most powerful mechanisms for sorting out children to assume different roles in the occupational hierarchy (p. 13).

Justice in decision making cannot result from test information alone, but, as George (1962) pointed out, standardized tests do help to eliminate much subjectivity.

Intelligence

Brim *et al* (1969) stated that the number of intelligence tests given each year continues to rise. The importance accorded to the measurement of intelligence was well expressed by Hunt (1961) as follows:

> Intelligence has long been of central concern to those seeking to understand the behavior of living organisms and especially the behavior of man (p. 3).

There have been many attempts to define intelligence. Anstey (1966) said that the various definitions of intelligence fall into three groups. The first group, which was labeled "biological," includes definitions like "capacity to learn" and "capacity to adapt to the environment." Examples of the second group, or "psychological" definitions, are "general mental efficiency" and "the capacity for abstract reasoning." The third type of definition of intelligence is considered "operational" and a popular example is "the sum of what intelligence tests measure."

Various theories of intelligence have also been developed and grouped. Brown (1970) said that the simplest approach is to postulate a unitary ability. Spearman (1927) proposed a two factor theory of intelligence. The first factor is a general factor which is operative at all times and the second is a specific factor. The specific factors are not correlated. Brown (1970) claimed that the group-factor approach is the one which is most popular. Thorndike developed a group-factor theory which describes intelligence as a composite of a multiple of overlapping and interrelated abilities (Thorndike et al, 1927). Thurstone (1947) introduced the idea that intelligence consists of several primary factors, each relatively independent of the other. Another representative of the

multiple factor theory is Guilford (1956) who suggested that there may be as many as one hundred twenty factors that make up the intellect.

Wechsler (1958) wrote that:

> Although Binet himself on several occasions made attempts to delimit the term, his primary concern was not with the definition but with the measurement or appraisal of intelligence, and this has been the main approach of physchologists since (p. 3).

But Wechsler (1958) went on to define intelligence as follows:

> Intelligence, operationally defined, is the aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal effectively with his environment (p. 7).

Jensen (1969) expressed his view in the following manner:

> Intelligence, like electricity is easier to measure than to define. And if the measurements bear some systematic relationships to their data, it means we can make meaningful statements about the phenomenon we are measuring. There is no point in arguing the question to which there is no answer, the question of what intelligence really is (p. 1).

Since we are able to measure intelligence and make meaningful statements regarding these measurements, it is not really necessary that there be agreement in defining intelligence.

Intelligence Tests

Most intelligence tests in use today are based on the "Stanford-Binet Intelligence Scale." Adams (1964) wrote that:

> For more than twenty years the "Stanford-Binet" was the standard measure of intelligence, the criterion with which all other intelligence tests, group and individual, were compared (p. 184).

Jensen (1969) stated that the first really useful intelligence test was the "Metrical Scale of Intelligence" divised in 1905 by Binet and Simon. This test has since undergone many revisions and today, in a form developed by Terman and Merrill, known as the "Stanford-Binet", it is generally considered the standard for the measurement of intelligence.

Another individual intelligence test comparable to the "Stanford-Binet" was the "Wechsler-Bellevue," published in 1939. This test has now been replaced by two scales: the "Wechsler Adult Intelligence Scale" and the "Wechsler Intelligence Scale for Children".

The first group intelligence test, the "Army Alpha," was developed in 1917 for use in the American army. Because of the need to test the illiterate and the non-English speaking soldiers, a second group test was developed. This test, the "Army-Beta", was basically a performance or

non-verbal type of group test. After the war these tests were released for civilian use. Several revisions were made and many new group intelligence tests were published and widely used. Anastasi (1968) expressed her view of this period as follows:

> The application of such intelligence tests far outran their technical improvements. That the tests were still crude instruments was often forgotten in the rush of gathering scores and drawing practical conclusions from the results (p. 12).

When these tests failed to meet their expectations a hostility towards tests developed which may have been harmful to the progress of intelligence testing (Anastasi, 1968).

Since this period, tests have been constantly revised and improved, and many new ones have been developed. They have proven very useful, although there are still limitations. Tyler (1963) said that intelligence tests measure a trait which is probably more limited than we suspect and they do not measure the human quality (human relations), special talents (art, music) or the adaptability of a student. She suggested that they measure the ability to deal with symbols, which is useful, but we must consider much more about people. Tyler (1963) summarized her view of intelligence tests as follows:

> With all their faults, intelligence tests are an indispensable tool in modern society. We use them to help us make decisions about the placement of individuals for school and work and to help us formulate educational and social policies. As research on intelligence tests has continued, they become sharper and more adaptable. But like all tools they require skillful handling and thorough knowledge of what they will and will not do (p. 54).

As long as educators recognize that intelligence tests do have limitations, both individual and group intelligence tests can be useful in our school systems today.

Group Tests and Individual Tests

Both group and individual tests are of value to educators. Morris (1961) found that group tests are in use in large numbers in schools. One of the reasons for the extensive use of group tests may be their ease of administration. The administrator does not need specific training other than to be able to follow the directions. Morris (1961) discovered that the majority of tests given in a school are administered by teachers. The scoring of group tests is often objective and can be done by a clerk or a machine. Group tests also take less time to administer. Often thirty or more students are tested in the same time that it would take to administer one individual test. Since time should be used efficiently this is a definite advantage of the group test. Cronbach and Gleser

(1965) expressed their view concerning testing time as follows:

> Among the chief problems of the decision maker is the efficient design of testing procedures. The time he can devote to testing is often severely limited; and how to use that time efficiently is a matter of concern (p. 86).

Another advantage of the group test is that it costs less per pupil than the individual test. Anastasi (1968) pointed out that group tests often provide better established norms because of the relative ease and rapidity of gathering data. The authors of group tests often use samples of 100,000 to 500,000 compared to samples of 2,000 to 4,000 used for norming individual tests.

The individual test, although not as widely used, has many advantages over the group test. It is a better test for children with special difficulties (mentally retarded, sensory handicapped) and for those who cannot perform adequately on a group test (Adams, 1964). In Chicago the WISC was used as a screening device along with an achievement test for youngsters with learning disabilities (Swanson and Jacobson, 1970). When the WISC was administered to youngsters with hearing difficulties it was found that the performance quotients were normally distributed and the mean and standard deviation did not differ significantly from the norms (Hine, 1970).

Also, the WISC has been used as a criterion for determining the underachiever (Teigland et al, 1966). In another study (Neville, 1965), it was concluded that poor readers in the fifth grade tended to make scores on group intelligence tests which were significantly lower than scores on individual intelligence tests.

When administering an individual intelligence test the examiner has a greater opportunity to observe individual behavior. Temporary conditions like illness, anxiety, fatigue, and so on, are more likely to be noticed. Also, the examiner, by observing, may learn more about the examinee than the test score.

The individual test usually has a greater variety of content which may help to keep the examinee's interest. The examiner also has the opportunity to motivate the examinee with encouragement or praise whenever necessary. The speed of testing can usually be varied to suit the student taking the test (Adams, 1964). The absence of competition with others and the opportunity for the examiner to establish rapport should help to stimulate the student to use maximum effort.

Generally the individual test is considered more accurate than the group test. Anastasi (1968) said that if the data are to be used for an important decision, it

is a good idea to give an individual test.

Testing Instruments Used

As indicated, most intelligence tests used today are based on the "Stanford-Binet Intelligence Scale." When discussing the tests used in this study (WISC and Otis-Lennon), reference is often made to the "Stanford-Binet."

The first Wechsler intelligence test, known as the "Wechsler-Bellevue" was published in 1939. Approximately ten years later the "Wechsler Intelligence Scale for Children" grew out of the Wechsler-Bellevue which was used with adolescents and adults. Like the WISC, the Otis-Lennon was also developed from an earlier edition. The Otis-Lennon is the fourth major edition of the Otis series which extends back to 1918 when the original Otis test was the first group intelligence test designed especially for schools.

Both the WISC and the Otis-Lennon attempt to measure general mental ability. The authors of both tests defined "intelligence" and then designed tests to measure this trait.

The Wechsler Intelligence Scale for Children

The WISC, designed for youngsters aged five to fifteen inclusive, received favorable comments from many writers. Horst (1966) stated that:

> Most good things that can be said about the "Stanford-Binet" apply equally well to the Wechsler Scales (p. 292).

Adams (1964) claimed that the WISC has a broader range of tasks than the "Stanford-Binet" because Wechsler selected items that represent different types of intellectual performance. The items of a given type, arranged in increasing order of difficulty make up subtests grouped into "Verbal" and "Performance".

There is substantial evidence to indicate that the WISC is highly reliable. Anastasi (1968) explained that a four-year follow-up study indicated that the WISC appears as stable as the "Stanford-Binet". Split-half reliability coefficients are reported in the WISC Manual for each subtest and for the verbal, performance, and full scale scores. They ranged from .92 to .95 for the full scale, from .88 to .96 for the verbal scale, and from .86 to .90 for the performance scale which indicates sufficient reliability for most testing purposes. Adams (1964) stated that the performance part is probably the most reliable performance test available. The subtest coefficients were not as high; some were near .50. The test manual cautions users against interpreting differences between subtest scores with reference to reliability coefficients. Some writers claim that the pattern analysis of subtest scores as a way of

distinguishing diagnostic categories has proved useful (Glasser and Zimmerman, 1968) while others feel that the rationale for interpreting subtest scores remains obscure (Cohen, 1959). Tyler (1963) concluded that the differences between verbal and performance scales are more meaningful than the discrepancies in the subtest scores.

There is no discussion of validity in the WISC manual, but Anastasi (1968) stated that concurrent validity coefficients between the WISC and other criteria, such as achievement tests and other measures of intelligence, cluster around .60. The verbal scale tends to correlate more highly than does the performance scale with such criteria. The verbal scale also correlates more highly with the "Stanford-Binet Intelligence Scale". Since the verbal and performance scales do not correlate that highly both parts are essential for a more sensitive measurement. Freeman (1962) concluded that the WISC Full Scale, Verbal Scale, and the "Stanford-Binet" have considerable communality of psychological functions being measured. Adams (1964) stated:

> The predictive validity of the Wechsler tests for academic criteria is somewhat reduced by the inclusion of the performance tests that are slightly less reliable and less closely related to academic success (p. 187).

Even with the inclusion of the performance part, the WISC is considered a good predictor of academic success

(Adams, 1964). Freeman (1962) suggested that more research on its predictive validity for educational purposes would prove useful.

The sample used for the standardization of the WISC was one hundred boys and one hundred girls at each of the eleven age levels. The children, representing eighty-five communities and eleven states, were all in school except fifty-five mental retardates. The distribution conformed closely to the United States census (1940) in terms of geographical area, urban-rural proportion, and parental occupation. The size and representativeness of this normative sample is considered very satisfactory but it is unfortunate that all youngsters were white. There is the possibility, however, that because no Negro or Mexican children were included, the sample might be more representative of the Canadian population than it otherwise would have been.

The intelligence scores are deviation scores (DIQs) obtained by comparing each subject's test performance with the scores earned by individuals in a single age group rather than a composite age group. The DIQs have a mean of 100 and a standard deviation of 15 whereas the subtest mean is 10 and the standard deviation is 3. Because the standard deviation of DIQs is identical from year to year a child's DIQ score should not vary unless his actual test

performance varies as compared to others his age (Wechsler, 1949).

Anastasi (1968) said that the WISC compares favorably with other individual tests in test construction procedures. The WISC is accompanied by a comprehensive manual with excellent directions for test administration and scoring. The directions are clear and complete and there is adequate basis for obtaining objectivity in scoring.

Some advantages of the WISC include the verbal/performance breakdown and the subtest format. The test items of each subtest are arranged in order of difficulty beginning with the easiest and progressing to the most difficult. When a youngster has gone as far as he is able in one subtest, he then starts at the beginning of the next subtest which helps prevent frustration. Another favorable feature is the great diversity and range of tasks which keeps motivation high. Other strong points are the manner of construction, standardization procedures and reliability which are comparable to the "Stanford-Binet" (Glasser and Zimmerman, 1968). The WISC is easy to administer and score, and it has proven to be an adequate predictor of academic success (Horst, 1966).

Some limitations of the WISC are that the standardization is too narrow in scope and there is only one form (Cohen, 1959). Cohen (1959) also claimed that the rationale

for interpreting subtest scores is not clear and that there are complicated wordings and emotionally laden items which may affect its validity.

Generally, the WISC is considered one of the best intelligence tests for youngsters aged eight to thirteen (Cohen, 1959). Many writers (Adams, 1964; Anastasi, 1969; Brown, 1970; Dubois, 1970; Glasser and Zimmerman, 1968; Horst, 1966; Tyler, 1963) expressed the same view as Mehrens and Lehmann (1969), who claimed that the WISC is a good test instrument.

The Otis-Lennon Mental Ability Test

The Otis-Lennon, designed for youngsters from kindergarten through Grade twelve, has received favorable comments from a number of reviewers (Grotelueschen, 1969; Smith, 1969). This test, of which there are two forms, was based on the Hierarchial Theory of Human Abilities formulated by Vernon (1960). This framework proved satisfactory because Otis and Lennon decided the test would measure general intellectual activity. Spearman's "G" factor is placed at the top of this hierarchy, followed by two general areas—"verbal-educational" and "practical mechanical". The broad "verbal-educational" group is divided into specific factors placed at the bottom of the hierarchy. It is in the "verbal-educational" portion of the hierarchial structure that the Otis-

Lennon was designed to function with the specific factors making up the content.

At each of the six levels of the Otis-Lennon, two forms were developed. Each form includes items balanced with respect to content, difficulty, and discriminating power (Grotelueschen, 1969; Smith, 1969).

The reliability compares favorably with other well constructed tests. Alternate forms, split-half and test-retest reliability coefficients were obtained to establish the precision of the Otis-Lennon as a measuring instrument in various conditions (Smith, 1969).

Otis and Lennon (1969) compared the results of their test to external-criterion measures to test for predictive validity. Most of the correlation coefficients between the Otis-Lennon and various achievement tests were in the .50 to .75 range. Correlations between the Otis-Lennon and teacher grades primarily ranged from .30 to .70 (Otis-Lennon, 1969).

Much care and planning went into choosing the sample which Groteluesche n (1969) claimed is representative of the United States kindergarten through Grade twelve school population. The standardization was stratified according to size and type of school system, socio-economic composition, and geographic region. Every state and over one

hundred school systems including public, Catholic, and private systems were presented. Nearly 4 per cent of the United States population was tested but youngsters from special classes for brain-injured, mentally retarded, and emotionally disturbed were not included. The scores can be converted to DIQs with a mean of 100 and a standard deviation of 16. The examiner can also obtain percentile ranks by age and grade, stanines by age and grade, and mental age equivalents.

The manual for administration includes an excellent discussion of the use of norms and interpretation of test scores. The directions for administering and scoring are clear and complete and the time limits are ample. Besides the administrator's manual, Otis and Lennon published a detailed technical manual which is relatively easy to understand and is informative. The test booklet itself is well constructed, well organized and well designed. The test battery and accessories for each level are of the same color which is a useful and attractive feature.

The authors, Otis and Lennon (1967) stated:

> The various levels comprising the Otis-Lennon Mental Ability Test series have been designed to provide a comprehensive, carefully articulated assessment of the general mental ability or scholastic aptitude of pupils in American schools (p.4).

After reviewing the literature concerning the Otis-Lennon Mental Ability tests it appears that the Otis-Lennon does provide a comprehensive, carefully articulated measurement for youngsters in kindergarten through Grade twelve. The Otis-Lennon, like the WISC, should prove very useful in the school setting.

HYPOTHESES

It would thus appear that both the WISC and Otis-Lennon are good test instruments for most students. But as Adams (1964) indicated, youngsters with special difficulties seem to have the most difficult time functioning on a group test. Thus students with adjustment problems may also have some difficulty functioning on a group test. They would probably be more easily distracted than they would during an individual test. As indicated in the literature, the WISC subtest format and the great diversity and range of tasks help prevent frustration and keep the student motivated (Glasser and Zimmerman, 1968).

Since the WISC has proven useful for youngsters with learning disabilities (Swanson and Jacobson, 1970), the sensory handicapped (Adams, 1964; Hine, 1970), underachievers (Teigland et al, 1966), poor readers (Neville, 1965) and the mentally retarded (Adams, 1964), the author

expected that the youngsters with adjustment difficulties might score higher on the WISC than they would on the Otis-Lennon. Hence, the following hypotheses were formulated.

Hypothesis I. The students experiencing adjustment difficulties will achieve a higher DIQ mean on the WISC (Full Scale) than they will on the Otis-Lennon.

Hypothesis II. There will be no difference between the DIQ means of the WISC (Full Scale) and the Otis-Lennon for the control sample.

Hypothesis III. There will be no difference between the correlation coefficients of the WISC (Full Scale) and Otis-Lennon for the adjustment and control samples.

CHAPTER III

EXPERIMENTAL DESIGN

Procedure

This study was conducted in the schools of the Biggin Hill School District at Medley, Alberta. Medley is an urban center of approximately 8,000 people comprised mainly of Canadian Forces personnel. The youngsters of this "service community" seem to have at least as many adjustment difficulties as the youngsters of any other community. Three of the four schools at Medley have counseling services which the students use freely. For the purpose of this study a youngster was considered to have an "adjustment" problem of some nature if he had visited a counselor at least four times during the 1971-72 school term.

To limit the study to one level of the Otis-Lennon series, the youngsters tested were in Grades four, five and six. It is also at this age level that the WISC has proved to be very useful (Cohen, 1959).

A random sample of thirty students was chosen from the sixty-four students in Grades four, five and six who experienced adjustment difficulties. These sixty-four students had been seen at least four times during the 1971-72 school year by a school counselor and/or the Director

of Guidance Services for the Biggin Hill School District. To select the required thirty students, each of the sixty-four students was assigned a number. Then all the numbers were written on cardboard pieces which were placed in a box and thirty were drawn. The same procedure (Ferguson, 1966) was followed to select the students from the 702 students for the control sample. Two of the sixty-four students with adjustment difficulties were selected for the control sample, but neither of these were among the thirty selected for the sample of youngsters with adjustment difficulties. Both samples included youngsters aged nine to thirteen inclusive.

During the 1971-72 school year the subjects were given both the WISC and the Otis-Lennon. The WISC was administered to each youngster by one examiner. The examiner was an experienced counselor trained in individual testing.

If a youngster had received the Otis-Lennon from his classroom teacher during the 1971-72 school year this score was used. Most of the youngsters, however, had not had the Otis-Lennon so it was administered to them in groups by the counselor in the school, or by the Director of Guidance Services.

In the testing situations care was taken to control the physical conditions as much as possible. The required materials were organized and an effort made to have the

examinees in quiet rooms with proper heat, light, and ventilation. The examiners attempted to develop a friendly relationship with the students. Instructions in the test manual were closely followed in an attempt to assure standard testing procedures.

The tests were scored according to standard procedures and intelligence scores computed for the WISC (Verbal, Performance, and Full Scale) and for the Otis-Lennon. The intelligence scores of the adjustment sample were compared to the intelligence scores of the control sample.

In the process of comparing the intelligence scores, the measures of central tendency, dispersion, and correlation were calculated and compared. Table IX reports the median and range while Tables X and XI indicate the means, standard deviations, and correlations for the WISC (Verbal, Performance, and Full Scale) and for the Otis-Lennon. A t-test was applied to determine if the means of the results of the adjustment sample differed significantly at the .05 level from the means of the results of the control sample.

To test the three hypotheses, the level of significance adopted was .05. For Hypothesis I, the Cochran-Cox method (Ferguson, 1966) was used to determine if the adjustment sample achieved a higher DIQ mean on the WISC than on the Otis-Lennon. The same formula was utilized

again to test Hypothesis II (the difference between the DIQ means of the WISC and Otis-Lennon for the control sample). Since the WISC and Otis-Lennon have different variances the Cochran-Cox method was chosen. This test is designed to determine the significance between sample means where population variances are unequal. To test Hypothesis III (the difference between the correlations of the two tests for both samples), Fisher's Z_r transformation (Ferguson, 1966) was applied.

CHAPTER IV

PRESENTATION OF DATA AND RESULTS

This chapter includes the data pertinent to the relationship of the WISC (Full Scale, Verbal, Performance) and the Otis-Lennon for both groups of students. The DIQs were processed by the Educational Research Services, University of Alberta. The means, standard deviations and correlation coefficients were computed for the WISC and the Otis-Lennon scores.

DIQ Scores

The data obtained from the testing are summarized in Tables I to VIII inclusive.

TABLE I

FREQUENCY DISTRIBUTION OF WISC
FULL SCALE DIQ SCORES FOR THE ADJUSTMENT SAMPLE

Full Scale DIQ Score	Frequency Distribution	Cumulative Frequency	Cumulative Percentage
111-115	4	30	100.0
106-110	2	26	86.7
101-105	2	24	80.0
96-100	7	22	73.3
91-95	6	15	50.0
86-90	6	9	30.0
81-85	2	3	10.0
76-80	0	1	3.3
71-75	1	1	3.3

TABLE II

FREQUENCY DISTRIBUTION OF WISC
VERBAL DIQ SCORES FOR THE ADJUSTMENT SAMPLE

Verbal DIQ Score	Frequency Distribution	Cumulative Frequency	Cumulative Percentage
111-115	2	30	100.0
106-110	4	28	93.3
101-105	1	24	80.0
96-100	4	23	76.7
91-95	7	19	63.3
86-90	6	12	40.0
81-85	5	6	20.0
76-80	1	1	3.3

TABLE III

FREQUENCY DISTRIBUTION OF WISC
PERFORMANCE DIQ SCORES FOR THE ADJUSTMENT SAMPLE

Performance DIQ Score	Frequency Distribution	Cumulative Frequency	Cumulative Percentage
121-125	1	30	100.0
116-120	0	29	96.7
111-115	4	29	96.7
106-110	4	25	83.3
101-105	5	21	70.0
96-100	5	16	53.3
91-95	4	11	36.7
86-90	4	7	23.3
81-85	1	3	10.0
76-80	1	2	6.7
71-75	1	1	3.3

TABLE IV

FREQUENCY DISTRIBUTION OF OTIS-LENNON DIQ SCORES FOR THE ADJUSTMENT SAMPLE

Otis-Lennon DIQ Score	Frequency Distribution	Cumulative Frequency	Cumulative Percentage
111-115	1	30	100.0
106-110	1	29	96.7
101-105	3	28	93.3
96-100	5	25	83.3
91-95	8	20	66.7
86-90	5	12	40.0
81-85	4	7	23.3
76-80	3	3	3.3

TABLE V

FREQUENCY DISTRIBUTION OF WISC FULL SCALE DIQ SCORES FOR THE CONTROL SAMPLE

Full Scale DIQ Score	Frequency Distribution	Cumulative Frequency	Cumulative Percentage
126-130	2	30	100.0
121-125	2	28	93.3
116-120	4	26	86.7
111-115	5	22	73.3
106-110	4	17	56.7
101-105	8	13	43.3
96-100	2	5	16.7
91-95	1	3	9.0
86-90	1	2	6.7
81-85	1	1	3.3

TABLE VI

FREQUENCY DISTRIBUTION OF WISC
VERBAL DIQ SCORES FOR THE CONTROL SAMPLE

Verbal DIQ Score	Frequency Distribution	Cumulative Frequency	Cumulative Percentage
121-125	2	30	100.0
116-120	2	28	93.3
111-115	2	26	86.7
106-110	7	24	80.0
101-105	5	17	56.7
96-100	7	12	40.0
91-95	3	5	16.7
86-90	1	2	6.7
81-85	1	1	3.3

TABLE VII

FREQUENCY DISTRIBUTION OF WISC
PERFORMANCE DIQ SCORES FOR THE CONTROL SAMPLE

Performance DIQ Score	Frequency Distribution	Cumulative Frequency	Cumulative Percentage
136-140	1	30	100.0
131-135	2	29	96.7
126-130	3	27	90.0
121-125	3	24	80.0
116-120	3	21	70.0
111-115	4	18	60.0
106-110	7	14	46.7
101-105	0	7	23.3
96-100	4	7	23.3
91-95	2	3	10.0
86-90	0	1	3.3
81-85	1	1	3.3

TABLE VIII

FREQUENCY DISTRIBUTION OF OTIS-LENNON DIQ SCORES FOR THE CONTROL SAMPLE

Otis-Lennon DIQ Score	Frequency Distribution	Cumulative Frequency	Cumulative Percentage
141-145	1	30	100.0
136-140	0	29	96.7
131-135	0	29	96.7
126-130	1	29	96.7
121-125	1	28	93.3
116-120	4	27	90.0
111-115	5	23	76.7
106-110	7	18	60.0
101-105	5	11	36.7
96-100	2	6	20.0
91-95	3	4	13.3
86-90	0	1	3.3
81-85	1	1	3.3

The tables display the frequency distribution, cumulative frequency and cumulative percentage of the DIQs for both tests. The results shown in Tables I to VIII indicate that the adjustment sample scored lower on both tests than did the control sample. In the adjustment sample, 73.3% of the youngsters obtained WISC Full Scale scores of 100 or less compared to 16.7% for the control sample. For the Otis-Lennon, 83.3% of the DIQs were 100 or lower for the adjustment sample compared to 20% for the control sample.

Median and Range

The median and range of each test for both samples are reported in Table IX.

TABLE IX

MEDIAN AND RANGE OF DIQs FOR WISC AND OTIS-LENNON FOR BOTH SAMPLES

		Adjustment Sample		Control Sample	
		Median	Range	Median	Range
WISC	Full Scale	95.5	41	108	47
	Verbal	91.5	39	104	36
	Performance	100.0	53	111	55
Otis-Lennon		92.5	33	109	56

As indicated, the medians of the adjustment sample were lower than those of the control sample, for both tests. The WISC (Full Scale) medians were 95.5 and 108 for the adjustment and control samples respectively. Similarly for the Otis-Lennon, the median for the adjustment sample was 92.5 compared to 109 for the control sample.

Mean and Standard Deviation

Table X includes the means and standard deviations of the DIQs for the WISC and Otis-Lennon. Like the medians, the means of the control sample were higher than those of the adjustment sample for both tests.

TABLE X

MEANS AND STANDARD DEVIATIONS OF DIQs FOR THE WISC AND OTIS-LENNON FOR BOTH SAMPLES

	Adjustment Sample		Control Sample	
	Mean	Standard Deviation	Mean	Standard Deviation
WISC				
Full Scale	96.333	9.789	108.633	11.072
Verbal	94.000	10.272	103.700	9.143
Performance	99.367	11.118	112.900	13.215
Otis-Lennon	92.567	8.144	109.000	10.929

The means of the adjustment sample were 96.333 (WISC) and 92.567 (Otis-Lennon) compared to the means of the control sample which were 108.633 and 109.000 for the WISC and Otis-

Lennon respectively. The t-test was used to determine if there was a significant difference in the sample means for both tests. Since only two groups were being compared the t-test was used rather than analysis of variance. Courts (1966) explained that "Although either method is applicable, the t-ratio is easier to compute when comparing only two groups." (p. 280). For the WISC the t-ratio calculated was:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}}}$$

$$= \frac{108.633 - 96.333}{\sqrt{\frac{9.789^2}{30} + \frac{11.072^2}{30}}}$$

$$= 2.698 \qquad df = 58$$

For a two-tailed t-test with 58 degrees of freedom, a t-value equal to 2.000 is required for significance at the .05 level. Since the score obtained was 2.698, the WISC (Full Scale) means for both samples were significantly different at the .05 level.

Comparing the Otis-Lennon means for both samples, the t-ratio obtained was:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}}}$$

$$= \frac{109.000 - 92.567}{\sqrt{\frac{10.929^2}{30} + \frac{8.144^2}{30}}}$$

$$= 6.604 \qquad df = 58$$

The t-value required for significance at the .05 level is 2.000. Therefore the Otis-Lennon means for the two samples also differed significantly at the .05 level.

The adjustment sample received a mean of 96.333 on the WISC and 92.567 on the Otis-Lennon. To test Hypothesis I (The students experiencing adjustment difficulties will achieve a higher DIQ mean on the WISC (Full Scale) than they will on the Otis-Lennon.), the t-value obtained using the Cochran-Cox method was:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{x}_1 - \bar{x}_2}}$$

$$= \frac{96.333 - 92.567}{2.325}$$

$$= 1.620 \qquad df = 29$$

The critical t-ratio required for significance at the .05 level for a one-tailed test with 29 degrees of freedom is 1.699. Since the t-ratio obtained was 1.620, the difference between the two test means for the adjustment sample was not significant at the .05 level. Therefore Hypothesis I was rejected. For the adjustment sample, the WISC mean was not significantly higher than the Otis-Lennon mean at the .05 level.

The control sample obtained means of 108.633 and 109.000 on the WISC and Otis-Lennon respectively. To test Hypothesis II (There will be no difference between the DIQ means of the WISC (Full Scale) and the Otis-Lennon for the control sample.), the t-score was calculated using the Cochran-Cox method.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{x}_1 - \bar{x}_2}}$$

$$= \frac{108.633 - 109.000}{2.840}$$

$$= -.129 \qquad df = 29$$

The t-score required for significance at the .05 level for the two-tailed test with 29 degrees of freedom is 2.045. Since the t-score obtained was only -.129, the difference was not significant at the .05 level. Therefore Hypothesis II was accepted and it was concluded that there was no significant difference, at the .05 level, between the DIQ means of the WISC and Otis-Lennon for the control sample.

Inspection of the means also indicated that both samples scored higher on the Performance part of the WISC than on the Verbal part.

Correlations

Table XI displays the correlation coefficients for the WISC (Full Scale, Verbal, Performance) and the Otis-Lennon. The WISC Full Scale and Verbal Scale scores correlated more highly with the Otis-Lennon than did the Performance Scale score. For the adjustment sample the Full Scale score correlated most highly with the Otis-Lennon and for

the control sample the Verbal score correlated the highest with the Otis-Lennon.

TABLE XI

CORRELATIONS BETWEEN WISC (FULL SCALE, VERBAL, PERFORMANCE) AND OTIS-LENNON DIQs FOR BOTH SAMPLES

Adjustment Sample

		WISC			Otis-Lennon
		F.S.	V.	P.	
WISC	Full Scale	1.000	.813	.788	.574
	Verbal	.813	1.000	.362	.551
	Performance	.788	.362	1.000	.406
	Otis-Lennon	.574	.551	.406	1.000

Control Sample

		WISC			Otis-Lennon
		F.S.	V.	P.	
WISC	Full Scale	1.000	.867	.889	.728
	Verbal	.867	1.000	.651	.828
	Performance	.889	.651	1.000	.478
	Otis-Lennon	.728	.828	.478	1.000

For both samples the lowest correlation coefficients were between the WISC Verbal and Performance scores. Generally, the correlations of the WISC scores with the Otis-Lennon appeared to be higher for the control sample than for the adjustment sample. The correlation coefficient of the WISC (Full Scale) and the Otis-Lennon for the adjustment sample was .574 compared to .728 for the control sample. To test Hypothesis III (There is no difference between the correlation coefficients of the WISC (Full Scale) and Otis-Lennon for the adjustment and control samples.), Fisher's Z_r transformation was used. The correlation coefficients (.728 and .574) were first converted to Z_r values (.925 and .654). Then using Fisher's formula the normal deviate (Z-score) computed was:

$$Z = \frac{Z_{r_1} - Z_{r_2}}{\sqrt{\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}}}$$

$$= \frac{.925 - .654}{\sqrt{\frac{1}{27} + \frac{1}{27}}}$$

$$= .996$$

The Z-score required for the two-tailed test at the .05 level is 1.96. Since the Z-score obtained was .996, the correlation

coefficients were not significantly different at the .05 level. Therefore Hypothesis III was accepted. It appears that the WISC (Full Scale) and Otis-Lennon correlate just as highly for the adjustment sample as they do for the control sample.

In a post hoc examination of the data it was noted that the correlation coefficients of the WISC Performance score and the Otis-Lennon were quite similar for the adjustment sample (.406) and the control sample (.478). This was not the case for the WISC Verbal and Otis-Lennon correlations which were .551 for the adjustment sample and .828 for the control sample. To test if there was a significant difference in the WISC Verbal and Otis-Lennon correlations for the two samples, Fisher's Z_r transformation was applied. At the .05 level, a Z-score greater than 1.96 was needed for significance. The Z value obtained was 2.065. This indicated that the correlation between the WISC Verbal and Otis-Lennon was significantly higher for the control sample than for the adjustment sample.

CHAPTER V

DISCUSSION

The writer wishes to point out that the discussion regarding the WISC and the Otis-Lennon is based on DIQ scores only. The other uses, advantages and limitations of these tests were not considered because the only data obtained in this study were the DIQs.

Findings and Implications

The analysis of the data did not indicate any significant differences, at the .05 level, between the WISC (Full Scale) and Otis-Lennon DIQs for each group. There was no difference at the .05 level of significance between the WISC (Full Scale) and Otis-Lennon means for the adjustment sample. This led to the rejection of Hypothesis I (The students experiencing adjustment difficulties will achieve a higher DIQ mean on the WISC (Full Scale) than they will on the Otis-Lennon.). Likewise, there was no significant difference at the .05 level between the WISC (Full Scale) and Otis-Lennon means for the control sample. As a result Hypothesis II was accepted. (There will be no difference between the DIQ means of the WISC (Full Scale) and the Otis-Lennon for the control sample.) Hypothesis III was also accepted because the correlation coefficients of the two tests for each sample were not significantly

different at the .05 level. (There will be no difference between the correlation coefficients of the WISC (Full Scale) and Otis-Lennon for the adjustment and control samples.)

The WISC (Full Scale) mean of the adjustment sample was close to being significantly higher than the Otis-Lennon mean at the .05 level. Because this difference was near significance at the .05 level, the author suggests the use of the WISC or another individual test for youngsters experiencing adjustment difficulties. The group test, however, appears to be satisfactory for the total population. Of course, if information about the student, other than the DIQ, is required, other factors will have to be considered when choosing the test. To insure that there is sufficient staff to administer the individual tests required, it might mean that the number of school psychologists needs to be increased in some school systems.

Since the WISC (Full Scale) mean was close to being significantly higher at the .05 level than the Otis-Lennon mean for the adjustment sample, further research involving these tests and youngsters with adjustment difficulties is required. On the Basis of this study it might prove worthwhile to compare a group test with an individual test using youngsters with "severe" learning difficulties. A

possibility would be to use youngsters who had seen a counselor at least ten times rather than four times. Perhaps an instrument for detecting adjustment difficulties would prove useful in a study involving these youngsters.

Ancillary Findings

A finding of particular interest was the significant difference at the .05 level, in the scores obtained by the adjustment sample and the control sample. The adjustment sample scored significantly lower on both tests than did the control sample. It seems that the youngsters who saw a counselor at least four times may have had a lower aptitude for academic success than the other students. If this were the case, it might have been one factor which contributed to their adjustment difficulties.

Teachers and counselors might find it worthwhile to take a look at the goals set for these youngsters. Perhaps the expectations of the "schools" are not realistic. Possibly the difference in the two samples indicates a need for special classes and/or programs for youngsters with adjustment difficulties. Another possibility might be to have small classes and/or support staff to allow sufficient time for individualized instruction and attention.

Conclusion

In conclusion, the results of this study, while interesting to the writer and of possible interest to teachers, counselors and administrators, did not indicate any significant differences at the .05 level in the DIQs of the WISC (Full Scale) and Otis-Lennon. Generally the results of the Otis-Lennon were in close agreement with the results of the WISC (Full Scale). The WISC Verbal score, however, did not correlate as highly with the Otis-Lennon for the adjustment sample as it did for the control sample. Also, the WISC (Full Scale) mean was very close to being significantly higher than the Otis-Lennon mean for the adjustment sample. It was therefore suggested that the WISC or another individual test should be used when testing youngsters with adjustment difficulties. This is especially important since it is possible that the results of intelligence tests given to these youngsters might be used more frequently than the results of intelligence tests given to the average students.

B I B L I O G R A P H Y

BIBLIOGRAPHY

Adams, G.S. Measurement and Evaluation in Education, Psychology, and Guidance. New York: Holt, Rinehart and Winston, Inc., 1964.

Anastasi, A. Psychological Testing. (Third Edition), New York: The MacMillan Co., 1968.

Anderson, D. E. A Comparison of the Peabody Picture Vocabulary Test with the Wechsler Intelligence Scale for Children. The Journal of Educational Research, 1968, 62, 114-116.

Anstey, E. Psychological Tests. Belfast: The University Press, 1966.

Brim, O. G. Jr.; Glass, D.C.; Neulinger, J.; and Firestone, I. J. American Beliefs and Attitudes About Intelligence. New York: Russell Sage Foundation, 1969.

Brown, F. G. Principles of Education and Psychological Testing. Hinsdale, Illinois: The Dryden Press, Inc., 1970.

Cashen, V. M. A Different Technique to Group Testing. Personnel and Guidance Journal, 1962, 40, 728-729.

Chase, C. I. and Ludlow, H. G. Readings in Educational and Psychological Measurement. (ed.) Boston: Houghton Mifflin Co., 1966.

Courts, F. A. Psychological Statistics An Introduction. Homewood, Illinois: The Dorsey Press, 1966.

Cronbach, L. J. Essentials of Psychological Testing. (Third Edition) New York: Harper and Row, 1970.

Cronbach, L. J. and Gleser, G. C. Psychological Tests and Personnel Decisions. Urbana: University of Illinois Press, 1965.

DuBois, P. H. A History of Psychological Testing. Boston: Allyn and Bacon, Inc., 1970.

Ferguson, G. A. Statistical Analysis in Psychology and Education. New York: McGraw-Hill, Inc., 1966.

Ferinden, W. E. Jr. and Jacobson, S. Educational Interpretation of the WISC. Linden, New Jersey: Remediation Associates, 1969.

Freeman, F. S. Theory and Practice of Psychological Testing. New York: Holt, Rinehart and Winston, 1962.

George, H. V. Testing--Another Point of View. English Language Teaching, 1962, 16, 72-78.

Glasser, A. J. and Zimmerman, I. L. Clinical Interpretation of the WISC. New York: Grune and Stratton, Inc., 1968.

Goldman, L. Using Tests in Counselling. New York: Appleton-Century-Crofts, Inc., 1961.

Good, W. R. Misconceptions About Intelligence Testing. The Education Digest, 1954, 20, 14-16.

Grotelueschen, A. Otis-Lennon Mental Ability Test. Journal of Educational Measurement, 1969, 6, 111-113.

Guilford, J. P. Fundamental Statistics in Psychology and Education. New York: McGraw-Hill, 1956.

Guilford, J. P. The Nature of Human Intelligence. New York: McGraw-Hill, 1967.

Hine, W. B. The Abilities of Partially Hearing Children. British Journal of Educational Psychology. 1970, 40, 171-178.

Horst, P. Psychological Measurement and Prediction. California: Wadsworth Publishing Co., Inc., 1966.

Hunt, J. McV. Intelligence and Experience. New York: The Ronald Press Co., 1961.

Jensen, A. R. How Much Can We Boost I.Q. and Scholastic Achievement? Harvard Educational Review, 1969, 39, 1-123.

Knight, R. Intelligence and Intelligence Tests. London: Methuen and Co., Ltd., 1943.

Magary, J. F. School Psychological Services. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1967.

McReynolds, P. (ed.) Advances in Psychological Assessment. California: Science and Behavior Books, Inc., 1968.

Mehrens, W. A. and Lehmann, I. J. Standardized Tests in Education. New York: Holt, Rinehart, Winston, Inc., 1969.

Morris, J. R. To Test or Not To Test. National Association of the Secondary School Principals' Bulletin, 1961, 45, 112-117.

Neville, D. The Relationship Between Reading Skills and Intelligence Test Scores. The Reading Teacher, 1965, 18, 257-262.

Noll, V. H. Introduction to Educational Measurement. Boston: Houghton Mifflin, 1965.

Otis, A. S. and Lennon, R. T. Otis-Lennon Mental Ability Test Technical Handbook. New York: Harcourt, Brace and World, Inc., 1969.

Payne, D. A. and McMorris, R. F. Educational and Psychological Measurement. Waltham, Mass.: Blaisdell Publishing Co., 1967.

Remmers, H. H.: Gage, N. L.: Rummel, J. F. A Practical Introduction to Measurement and Evaluation. New York: Harper and Brothers, 1960.

Smith, A. E. Test Reviews. Journal of Counseling Psychology, 1969, 17, 91-92.

Spearman, C. The Abilities of Man. New York: The MacMillan Co., 1927.

Stodola, Q. and Stordahl, L. Basic Educational Tests and Measurement. Chicago: Science Research Associates, Inc., 1967.

Swanson, M. S. and Jacobson, A. Evaluation of the S.I.T. for Screening Children with Learning Disabilities. Journal of Learning Disabilities, 1970, 3, 318-320.

Teigland, J. J.; Winker, R. C.; Munger, P. F.; Kranzler, G.D. Some Concomitants of Underachievement at the Elementary School Level. Personnel and Guidance Journal, 1966, 44, 950-955.

Terman, L. M. and Merrill, M. A. Measuring Intelligence. Boston: Houghton Mifflin Co., 1937.

Thorndike, E. L.; Bregman, E. O.; Cobb, M. V.; Woodyard, E. The Measurement of Intelligence. New York: Teachers' College, Columbia University, 1927.

Thorndike, R. L. and Hagen, E. Measurement and Evaluation in Psychology and Education. (Second Edition) New York: John Wiley and Sons, Inc., 1961.

Thurstone, L. L. Multiple-Factor Analysis. Chicago: University of Chicago Press, 1947.

Travers, M. W. Educational Measurement. New York: The MacMillan Co., 1955.

Tyler, L. E. Tests and Measurement. New Jersey: Prentice-Hall, Inc., 1963.

Vernon, P. E. The Measurement of Abilities. London: University of London Press, Ltd., 1956.

Wechsler, D. The Measurement and Appraisal of Adult Intelligence. Baltimore: The Williams and Wilkins Co., 1958.

Wechsler, D. Wechsler Intelligence Scale for Children Manual. New York: The Psychological Corporation, 1949.

Whitla, D. K. Handbook of Measurement and Assessment in Behavioral Sciences. Cambridge, Mass.: Adisno-Wesley, 1967.

A P P E N D I C E S

Appendix A

Appendix B

Appendix C

Appendix D

A P P E N D I X A

APPENDIX A

Results of a Telephone Survey conducted in January, 1972.

The principals and/or counselors of nine schools in the Biggin Hill and Bonnyville schools systems were asked the following questions:

Question I: Are the intelligence test scores recorded in the cumulative records usually the results of individual tests or group tests?

Reply:

Individual	0
Group	9

Question II: Are the intelligence test scores recorded in the cumulative records used more frequently by teachers, counselors, and administrators for students with adjustment problems or for the average students?

Reply:

"Adjustment"	9
"Average"	0

A P P E N D I X B

APPENDIX B

Letter to Harcourt Brace Javanovick, Inc.

Box 1338, Medley,
Alberta, Canada.
January 23, 1972.

Mr. J. Wakeman, Regional Manager,
Harcourt Brace Javanovick, Inc.,
Test Department, Polk and Geary,
San Francisco, California 94109.

Dear Mr. Wakeman:

Re: Otis-Lennon Mental Ability Test

I am working on a research project based on the "Otis-Lennon Mental Ability Test." Please forward any recent literature involving this test.

Thank you for your assistance.

Yours sincerely,

(Mrs.) C. Pernarowski

A P P E N D I X C

APPENDIX C

DIQs As Measured By The WISC And The Otis-Lennon For The Adjustment Sample

Student	WISC			Otis-Lennon
	Full Scale	Verbal	Performance	
1	115	115	111	100
2	115	104	124	103
3	112	106	115	85
4	111	114	106	111
5	106	110	100	106
6	106	110	100	95
7	105	110	99	96
8	103	94	113	103
9	99	96	101	95
10	99	92	107	99
11	99	91	107	85
12	99	87	113	92
13	98	97	99	88
14	98	94	103	95
15	96	97	94	96
16	95	97	93	102
17	95	84	108	99
18	93	85	104	79
19	93	84	104	88
20	92	91	94	92
21	91	84	101	88
22	88	92	85	93
23	88	90	89	80
24	88	89	89	93
25	88	87	92	89
26	87	86	90	85
27	87	86	90	83
28	85	91	80	78
29	85	76	99	91
30	74	81	71	88

APPENDIX C

DIQs As Measured By the WISC And The Otis-Lennon For The Control Sample

Student	WISC			Otis-Lennon
	Full Scale	Verbal	Performance	
1	129	120	133	124
2	126	110	139	108
3	125	121	125	127
4	123	116	127	119
5	120	121	114	141
6	120	111	127	110
7	119	106	129	111
8	118	113	121	118
9	115	108	133	100
10	112	110	111	113
11	112	101	121	111
12	111	104	117	110
13	111	103	118	109
14	109	108	108	117
15	108	106	108	113
16	108	96	120	106
17	107	105	107	116
18	105	100	110	104
19	104	106	100	114
20	104	97	111	105
21	104	97	110	107
22	104	94	115	102
23	103	97	108	109
24	103	97	108	103
25	101	104	99	99
26	98	97	99	95
27	96	94	100	94
28	93	95	93	105
29	89	89	92	95
30	82	85	84	85

A P P E N D I X D

APPENDIX D

Sample Letter to Principals at Schools Involved in this Study.

Box 1338,
Medley, Alberta.
September 1, 1972.

Mr. R. Safroniuk,
Principal,
MacKenzie School,
Medley, Alberta.

Dear Rudy:

May I express my sincere thanks to you and your staff for the co-operation received while collecting the data for my research project.

The results will be forwarded to your school in October, 1972.

Yours sincerely,

Chris Pernarowski,
Director Guidance Services.